

COMPUTER APPLICATIONS IN ONCOLOGY

Development and implementation of computer techniques for collection,
processing, and retrieval of medical information and for documentation
of external beam therapy

TORGIL R. MÖLLER

LUND 1977

COMPUTER APPLICATIONS IN ONCOLOGY

Development and implementation of computer techniques for collection,
processing, and retrieval of medical information and for documentation
of external beam therapy

AKADEMISK AVHANDLING

som med vederbörligt tillstånd av Medicinska fakulteten i Lund
för avläggande av medicine doktorsexamen kommer att
offentligen försvaras i Onkologiska klinikens föreläsningssal,
Lasarettet i Lund, lördagen den 7 maj 1977 kl. 9.15

av

TORGIL R. MÖLLER

med. lic., Ssk

LUND 1977

COMPUTER APPLICATIONS IN ONCOLOGY

Development and implementation of computer techniques for collection,
processing, and retrieval of medical information and for documentation
of external beam therapy

TORGIL R. MÖLLER

Department of Oncology, University Hospital, S-221 85 Lund, Sweden



This summary is based on the following publications:

- I Gustafsson T., Möller T.R., and Nordberg U.-B.:
BEKO - A system for documentation of external beam
treatment in radiotherapy.
Computer Programs in Biomedicine 3 (1973), 135-144.
- II Möller T.R., Johnsson J.- E., Sundström A., and
Gustafsson T.: Computer-based patient record and
follow-up system for radiotherapy.
In: Computer Applications in Radiation Oncology,
Proceedings of the Fifth International Conference on the
Use of Computers in Radiotherapy, Hanover 1974.
University Press of New England, Hanover 1976, p. 30-34.
- III Möller, T.R., Nordberg U.-B., Gustafsson T., Johnsson
J.-E., Landberg T.G., and Svahn-Tapper G.: Planning,
control, and documentation of external beam therapy.
Acta radiologica (1976) Supplementum 353.
- IV Möller T.R., and Gustafsson T.: An on-line method for the
acquisition of medical information for computer processing
as applied to radiotherapy.
Computer Programs in Biomedicine 7 (1977) In press.

These publications will be referred to in the text by their Roman numerals.

CONTENTS

INTRODUCTION	7
TECHNIQUES OF DATA COLLECTION	11
Abstract sheets	11
Data collection based on a structured radiotherapy record	12
Treatment record form	14
Discussion	16
CONTENTS OF THE DATA BASE	19
Personal data	19
Medical history	20
Diagnostic data	20
Treatment data	21
Follow-up data	23
Discussion	23
FUNCTIONS OF DATA RETRIEVAL	26
Main patient register file	26
"Statistical" file	27
Treatment documentation system	32
Discussion	32
VALIDITY CHECKS	35
Computer checking procedures	35
Manual checking procedures	36
Validation study of the present system	36
Discussion	38
DATA PROCESSING	41
Batch systems	41
Direct access routine	42
Local interactive system	42
Discussion	43
GENERAL SUMMARY AND CONCLUSIONS	46
ACKNOWLEDGEMENTS	49
APPENDIX	51
REFERENCES	64

ABBREVIATIONS

CRN	Civic registration number (Personal identification number for Swedish citizens).
ICD	International Classification of Diseases.
ICRU	International Commission on Radiation Units and Measurements.
MOTNAC	Manual of Tumor Nomenclature and Coding.
MUMPS	Massachusetts General Hospital Utility Multi-Programming System (Data base operating system).
SNOP	Systematized Nomenclature of Pathology.
TNM	Classification system for malignant tumours (T = tumour, N = regional nodes, M = metastases).
UICC	Union Internationale Contre le Cancer.
VDU	Visual display unit.

INTRODUCTION

The first application of computers in radiotherapy is attributed to TSIEN in 1955 with the calculation of absorbed dose distribution in external beam treatment planning. During subsequent years, the number of applications continuously increased and today computers may be utilized in most of the steps in the radiotherapy procedure, as described in Fig. 1.

Computers may increase the yield of information of procedures in diagnostic radiology and nuclear medicine. Computer-assisted methods to obtain cross-sectional topography necessary for treatment planning, such as computerized tomography, have been developed. Actual dose calculation is performed more rapidly by computer than by manual methods, and corrections for tissue heterogeneity may be included to give a more accurate dose distribution pattern. Quantitative information may be obtained to facilitate the selection of the optimal dose plan. In the daily treatment procedure, computers may provide all responsible personnel with information on treatment parameters, and various monitoring systems may assist in the set-up of patients (EDWARDS 1974, III).

Thus, computer technology may in different ways contribute to an increased accuracy in delivering radiation therapy. However, improvements in medical care must be assessed in terms of cure rates, morbidity rates and survival rates. A continuous evaluation of different treatment techniques for different series of patients, based on long-term follow-up, is thus necessary. A considerable amount of data has to be collected and analysed to enable such evaluations. By manual means, this is an almost impossible task, and therefore manual evaluations are usually confined to a limited number of problems. Here, computers may have a great impact on the possibilities of improving medical care due to their capacity of storage and speed in processing large amounts of data.

Electronic data processing of medical records was tried rather early in the advent of computers, and the amount of work spent on such projects is immense. Despite this, however, the number of systems in actual clinical use is limited (COLLEN 1974), chiefly due to the complexity of such systems with corresponding high running costs. With the rapid development of mini-computers during the last years, a number of departmental medical informa-

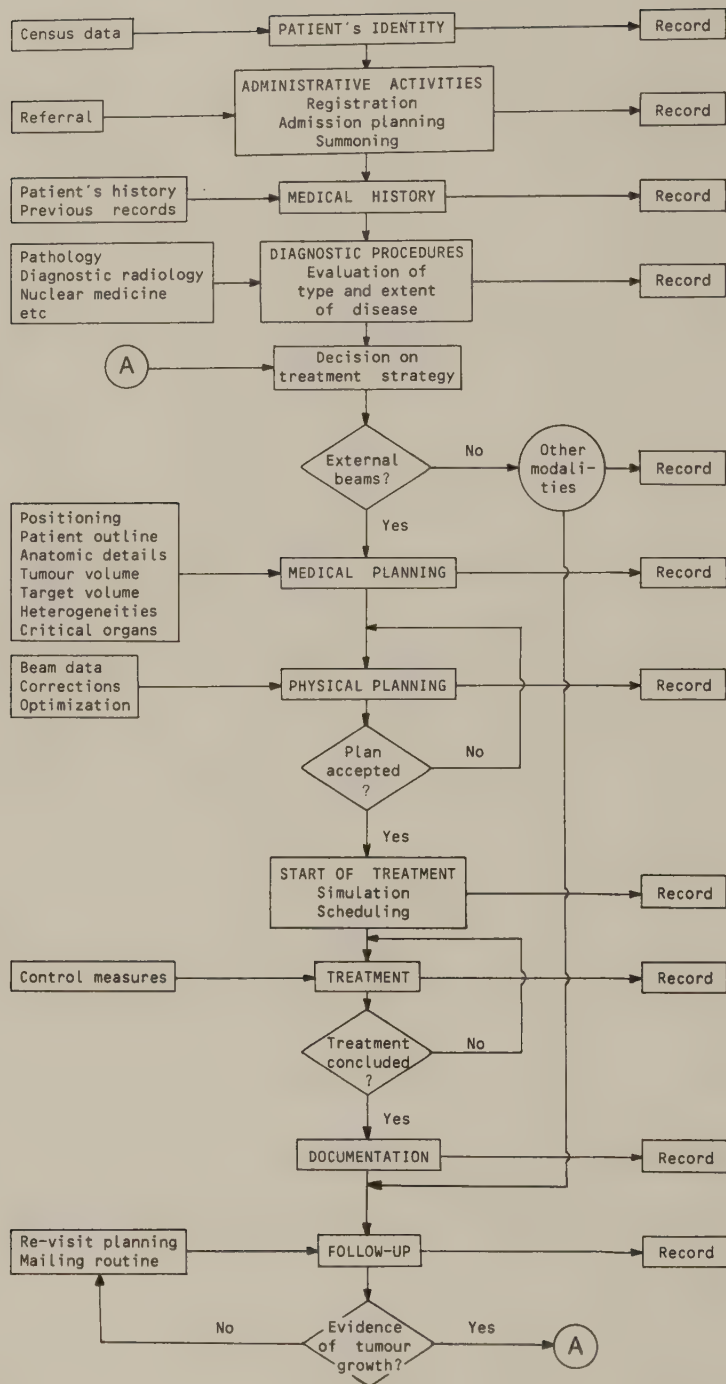


Fig. 1 Flow-chart describing the procedure of external beam therapy (modified from ICRU Report 24, 1976). The various steps, at which data recording may take place, are indicated. Similar flow-charts may be drawn for the other treatment modalities employed in oncological care.

tion systems, based on minicomputers, have been developed.

At the Department of Oncology in Lund, work was begun in 1967 on the design of a computer-based system for patient follow-up and treatment evaluation. This necessitated the systematization and definition of a number of treatment parameters and the definition of a data base. Data collection was started in 1968 on a limited scale, with data processing on a computer in a neighbouring town. With the installation at the Department in 1971 of a remote job entry terminal, linked to the Lund University Computer Centre and the implementation of a new data collection technique, a step was taken towards a more complete information system. Concurrently, systems for image processing in nuclear medicine and external beam treatment planning were developed and introduced into the clinical routine. Presently, with the installation of a local computer at the Department in 1975, all systems are being converted to local service.

The main objectives for starting the project of computerization of medical information in radiotherapy at this Department were:

- ☐ to keep an up-to-date register of all patients referred to the Department for reference purposes;
- ☐ to enable search and listings of groups of patients fulfilling certain criteria for further analysis of case note data;
- ☐ to provide annual statistics on various aspects of the patient population and treatments;
- ☐ to obtain survival rates, cure rates, and morbidity rates for different groups of patients selected by e.g. diagnosis, stage and treatment technique;
- ☐ to get automatic print-out of treatment reports;
- ☐ to automatize the mailing follow-up routine.

The primary intentions were not to create still another system for computerized record keeping but continuously to collect, store and process a limited number of well-defined data necessary to reach the above-mentioned goals.

The aim of this work was therefore:

- ☐ to define the contents of the data base;
- ☐ to develop a data collection system for the recording of selected information;
- ☐ to define a number of variables used in radiotherapy procedure;
- ☐ to develop a system for documentation of external beam therapy;
- ☐ to design and implement a number of data retrieval functions.

The programming and technical aspects of electronic data processing were considered to lie outside the scope of this thesis and will not be dealt with.

TECHNIQUES OF DATA COLLECTION

To get an estimation of the annual data collection volume, the numbers of new patients, treatments, and follow-up contacts at the Department of Oncology in Lund are given for 1975 in Table I. At each of these listed events, a certain number of variables should be recorded.

TABLE I. Numbers of patients and treatments 1975 at the Department of Oncology, University Hospital, Lund.

Patients' admissions	2 229
In-patient episodes	2 706
Visits to out-patient clinic	16 689
Follow-up contacts by mail	≈ 6 000
Treatment series, external beam	2 233
Treatments, external beam	34 829
"-", surface application	32
"-", intracavitary	502
"-", interstitial	9
"-", internal (metabolic)	289
Chemotherapy courses	405
Operations (gynecologic section)	125

ABSTRACT SHEETS

The information to be collected was initially organized in distinctive subgroups and divided between three separate sheets (MÖLLER 1970). These proformas are described in the Appendix (Figs. 9-13):

- A Patient identity, administrative data on referral and admission, medical history, diagnostic data, and surgery;
- B External beam therapy, surface application, intracavitary and interstitial therapy, internal (metabolic) therapy, chemotherapy, and hormonal therapy;
- C Status at follow-up, side effects of treatment, measures to be taken, and date and cause of death.

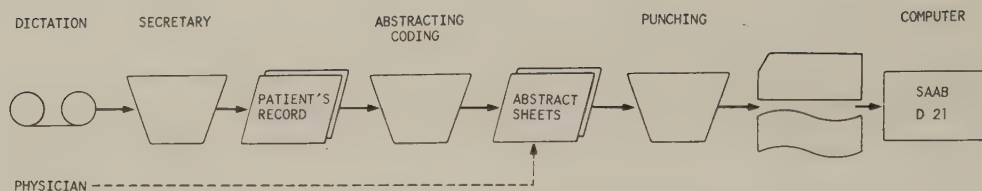


Fig. 2 1968: Data collection based on special abstracting and coding sheets. These are filled in by the secretary; the possibility of having the physician to complete the forms is only indicated, since this alternative rarely will function.

After completion, the proformas were collected and the data punched on paper tape for further processing as described on page 41. The estimated annual number of proformas was 2 500 for sheet A, 4 000 for sheet B, and about 20 000 for sheet C.

This system had a trial run for a period in 1968, but could not attain the intended level of functioning, chiefly due to the fact that filling in had to be done concurrently with ordinary record keeping and this duplicating of registration imposed too much extra workload on the medical staff. In reality, the proformas had to be completed at a later date with information extracted from the case notes (Fig. 2). The routine was therefore abandoned, except for the registration of identity and diagnosis. For the collection of the rest of the desired information other means had to be developed. Following the principle of collecting data as near to its source as possible, it was decided to collect medical information when typing the medical record, and data on external beam therapy at the treatment record level.

DATA COLLECTION BASED ON A STRUCTURED RADIOTHERAPY RECORD

A structured record, as described in IV, is a prerequisite of the current data collection techniques (II, IV).

Practically all medical information concerning each single patient is entered into his type-written case notes in chronological order. Excluded are the results of certain laboratory tests, analogue information, and pictorial information. Thus the record file includes the medical history, physical examinations, the results of diagnostic and therapeutic procedures such as microscopy, x-ray examinations, surgery, daily progress notes, and discharge summaries. Each piece of information is identified by a special heading, the complete set of which is given in IV. These headings are in consistency with

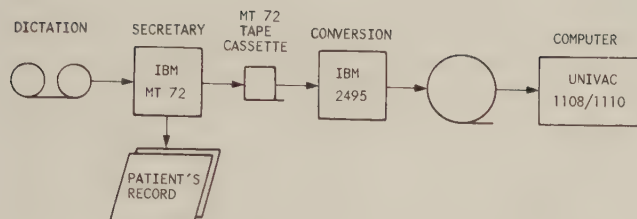


Fig. 3 1971: Data collection based on an off-line magnetic tape encoder. The patient's record is typed after the dictation according to the structure described. Information is recorded on a magnetic tape cassette as combinations of code+heading+string. These are later decoded in the central computer by a special set of programs.

suggested national standards (SPRI 1976). The information entered under each heading may be a value, a structured string, free text, or certain sub-headings.

The lay-out of the typewritten record is also described in IV. When typing the record according to this format on a special typewriter the headings and the typewritten text may be collected on a separate medium for further processing by computer.

Off-line system

The first version of this data collection system, as described in II, was based on a magnetic tape encoder (IBM MT 72). This is in principle an electric typewriter, equipped with two magnetic tape cassettes, called "master tape" and "slave tape". On the master tape, all headings, sub-headings and tabulator positions used in the typing of the record were stored. Certain headings were grouped together, each group or module relevant to a specific piece of information. The master tape was divided into thirty modules in order to cover all aspects of medical information. The internal structure of each module had to be very rigid and all agreed expressions that could be entered in certain structured strings were stored on the master tape. This necessitated quite a number of "START" and "SKIP" instructions.

Simultaneously as the heading was printed on the record sheet it was recorded on the slave tape. The typewritten text under the heading was subsequently also recorded on this tape. The cassette-tapes were later converted to a standard half-inch 9-track computer compatible magnetic tape, which updated the main register at regular intervals (Fig. 3). At updating, the relevant information was extracted, decoded, and added to the main register file (II).

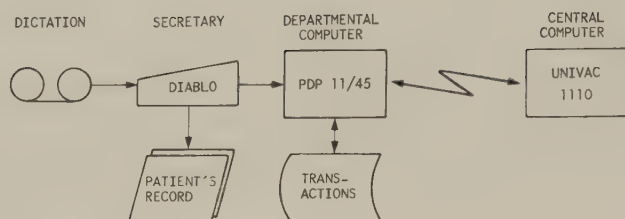


Fig. 4 1976: Data collection on-line. The record is typed on a typewriter terminal linked to a minicomputer. Recognition of headings and generation of internal codes are done automatically. The transactions are stored in a temporary disk file, and transferred to the central computer at regular intervals.

On-line system

Although the off-line system did function quite well during a number of years, the running costs, the lack of flexibility of the master tape lay-out, and increasing number of technical faults with increasing age of the equipment were annoying. Therefore it was decided to transform the data collection system into an even more user-oriented system on-line, as described in IV. The structure of the radiotherapy record was kept unchanged, but the record was typed on a typewriter terminal linked to a departmental minicomputer. With this arrangement it was possible to have automatic recognition of the headings, and they could be entered in an arbitrary order. In cases where a structured string had to be entered, this was also recognized automatically. The identification of a heading (or string) was confirmed by the machine printing the remaining characters, positioning the typewriter head, and shifting the colour of the ribbon. If the initial characters of a heading were not in conformity with the accepted words, no echo was given and thus no wrong character was printed.

The transactions were stored cumulatively on a disk. These transactions could form the input either to the main register at the central computer or to the data base on the local minicomputer (Fig. 4).

Some of the properties of the three data collection systems are summarized in Table II.

TREATMENT RECORD FORM

For external beam therapy, a special treatment record form was developed (I, III). At the conclusion of treatment, a number of variables (listed in I)

TABLE II. Comparison of three typewriter based data collection systems.

	Abstract sheets	Structured record, off-line	Structured record, on-line
Equipment	Eichner Tronictyper (punch + reader)	IBM MT72	NOKIA 30 KSR (Diablo)
Recording medium	Paper tape off-line	Magtape cassette off-line	Disk on-line
Flexibility in sequence of entry	Varying but rather good	No flexibility within the same master tape module, but modules in arbitrary order.	Total flexibility
Validity checking	None	Signal when data recorded on slave tape, else none	CRN, headings and sub-headings; extensive development possible
Correction facilities	Copying entire paper tape	Corrections entered on the same magtape cassette	No echo if wrong char. in heading; backspace when typing; text editor
Hard copy lay-out	Mostly codes	Medical record sheet	Medical record sheet
Technical reliability	Poor: frequent breakdowns ($\approx$ once weekly)	Rather good: infrequent breakdowns ($\approx$ once monthly) conversion trouble with old magtape cassettes	Very good: no breakdown so far
Printing speed	10 char/s	10 char/s	30 char/s
Noise	Much	Moderate	Little
Operator's acceptance	?	Quite acceptable	Very positive
Costs	Purchase: Skr 25 000 (1969), paper tape	Leasing: Skr 20 000/year + conversion costs Skr 5 000/year, cassettes Skr 100/each (purchase)	Purchase: Skr 20 000 - 30 000; requires mini-computer

TABLE III. Comparison of two methods of data collection from the treatment record form.

	Card punching	Dialogue mode, on-line
Equipment	IBM 29 card punch	NOKIA 30 KSR (Diablo) typewriter terminal or VDU
Recording medium	Punched cards	Disk on-line
Recording characteristics	Duplication of "header" on each card; all variables must be re-coded	Some variables generated automatically; questions branched according to answers
Validity checking	None	CRN; extensive development possible
Errors: typing error rate	1 in 2 record forms (estimated) (1 in 120 variables)	1 in 5 record forms (1 in 300 variables)
- - : most common error	Column displacement: will affect all variables following erroneous entry on same card; duplicating may introduce systematic errors	Row displacement: will only affect involved variables
Correction facilities	Punch new card	Text editor
Recording speed	3 - 4 record forms/h	10 record forms/h
Noise	Much	Little or no
Operator's acceptance	Negative	Positive

were collected for further processing. Thus, this record form replaced the proforma for the recording of external beam therapy described previously.

At the start, data collection was performed by key-punching on cards. This type of data collection, applied to such a complex document as the treatment record form was, however, very prone to errors, especially column displacement. With the installation of the local computer, data collection was therefore transferred into an interactive dialogue system on a typewriter terminal or a VDU. A comparison of these two types of data collection is given in Table III.

DISCUSSION

The ideal data input system should fulfill the following requirements as defined by BROSS (1973): "The data input system should not disturb the institutional environment. The system should be versatile enough to be able to handle almost all of the information items in a record. The system should be able to "evolve" to meet the need of its users. The system should make minimal demands on the users and on the personnel maintaining it. There should be prompt, effective, quality-control of the input. The system should be reliable. The cost of the system should not be excessive." Surveys of different data collection systems are given by e.g. DEELEY (1972) and BRINCKER (1977), and bibliographies by McLEAN et coll. (1975) and KÖHLER et coll. (1976).

The most common of all data collection and input systems is the use of abstract sheets or proformas, as described by KAHN & MOSKON (1965), ELLIS (1968, 1970), SNELLING (1970), MINET (1970), BRINCKER (1971), McAINSH et coll. (1971), EDWARDS (1972, 1974), DUTREIX et coll. (1972), STRÜTER & HALAMA (1974), WHEELER et coll. (1976) among others. Such proformas have definite advantages; the information may be structured logically, they may function as an "aide memoire" to the physician, and it is easy to ensure that the collection of data is complete. One of the disadvantages is that they usually have to be filled in by the physician, thus increasing his work-load, or by a coding clerk or secretary. Although coding clerks may be quite efficient, the quality of data has been shown to deteriorate when passing through this intermediate step. Still, information has to pass a punch operator, where errors also may be introduced.

If the proforma is designed in such a way that it can be read by a magnetic or optical document reader (ELLIS 1970, GROS 1970, WHEELER 1976), this last intermediate step may be omitted. Optical character readers have also been used as input devices for documents, typed on a typewriter with a

special font (BRUCE et coll. 1970, BRUCE 1971).

Typewriters, equipped with paper tape, magnetic tape (KOREIN 1970, ROBINSON 1970) or magnetic cards (BROSS 1973) have been used as data collection devices, mostly for the collection of free or narrative text. The two latter types offer great potential for data collection, but certain limitations as regards flexibility may be noted as described above.

Ideally, data should be recorded at the site of origin and at the time when generated. To fulfill this demand it is usually necessary to make the physician either complete a proforma or operate an on-line terminal such as a VDU. For this to be feasible, the physician has to be relieved of the task of ordinary record keeping, if not a double work-load is to be put on him. Thus it has been suggested (e.g. MacKINNON 1970) that the proformas should replace the ordinary record. However, it is doubtful if a proforma, carefully designed as it may be, can fully be a substitute for the ordinary patient's record with its unlimited flexibility. If the entire record is computerized, the need for print-outs may put a heavy load on the computer, and sufficient accessibility requires a reasonable number of terminals to be available, demanding a high operational performance of the computer with practically no down-time.

For the current type of structured record the sources of information are either the physician's dictation or a written document. In both cases, fixation of data occurs when the secretary types the record, and this point of time may be chosen as the best one for the collection of the various types of data.

The physicians at this Department have since long been accustomed to dictate their observations and to edit information obtained from various sources. The experience of daily routine during many years has shown, that it is easy to organize the dictate according to the proposed structure of the record, and this also facilitates the secretary's work when typing. To type the record on a standard typewriter or a typewriter linked to a computer is a matter of cost, since the attitude of the secretaries to this type of data collection has been positive.

The need for organizing the case notes have been clearly stressed by Weed when designing the "problem oriented record" (WEED 1968, WAKEFIELD & YARNALL 1976). Even if the record is kept at a single department and not common for a hospital, the application of Weed's ideas and concepts may be advantageous. At a department of oncology, it is natural to regard a malignant disease as a defined problem, and in the case the patient suffers from more than one tumour, each malignancy constitutes a separate problem. Thus, it must be possible to relate diagnostic procedures, treatments and clinical observations to the pertinent diagnosis / problem. In the structured record as described in IV, each new entry is preceded by a Diagnosis No..

If the information is relevant to all malignancies in a patient with multiple tumours, this entry may be left blank.

Data collection according to the method described has the definite advantage of producing a hard-copy at the same time as the information is recorded on disk in the computer. The hard-copy, which is in fact the patient's updated case notes, also serves as a check that data recording has taken place.

When updating the main register with new transactions, recorded either on the slave tape or on the disk, special programs perform further extraction and decoding of data to be included in the data base. Since only certain strings having a specified structure are decoded, any processing of free text, as described by e.g. GELL & SCHWARTZ (1974), OKUBO et coll. (1975), FENICHEL & BARNETT (1976) and SCUDIERO & WONG (1976) was not included. In our opinion, the complexity of such systems did not favour this approach to data extraction. However, analysis of the free text describing the clinical and microscopic diagnoses, such as described by GREENWOOD (1972), may be a valuable step in the checking procedures. Still, by the use of structured strings the number of codes have been kept to a minimum (ICD, SNOP, TNM, and codes for parish and referring department).

CONTENTS OF THE DATA BASE

The data base described here was intended for the evaluation of treatment delivered at a Department of Radiotherapy for any type of malignant disease. Originally defined for the abstract sheets described previously, it is still valid for the current data collection techniques, but certain parts of the data base have been subject to revision to suit the changing pattern of activity at the Department.

Since a detailed listing of the data base is given in the Appendix in conjunction with the appropriate proformas (Table VIII), only comments on some of the items of the data base will be given here.

PERSONAL DATA

The problem of unequivocal identification of a patient is complex. In the Scandinavian countries, however, this problem is simplified by the use of a special civic registration number (CRN). This is a ten-digit number, consisting of date of birth, a regional sequence number indicating place of birth and sex (odd numbers for males and even numbers for females), and a check digit. Record linkage is in practically all situations based on this personal number, and therefore all efforts must be made to ascertain that the correct CRN is recorded for the patient.

If not known to the patient himself, the complete CRN can easily be obtained from e.g. a population register on microfiche at the registration office of the department.

The registration of social and demographic data was not thought pertinent to a data base intended for the evaluation of radiotherapy. However, if the data base forms part of a hospital or regional tumour registry, a detailed history on past and present occupations, hobbies, and exposures to different agents might be included. The same comments apply to information on religion, ethnic group, and skin colour (WHO 1976). With increasing immigration the importance of such data may be enhanced.

MEDICAL HISTORY

Information on medical history is limited and mainly intended to distinguish patients with a positive family history of malignancy, patients with a previous or concurrent malignancy, and patients treated prior to admission.

For specific diseases, a more comprehensive medical history may be desirable, but such is preferably included in the disease specific parts of the oncology data base.

DIAGNOSTIC DATA

The necessary elements for the description of a tumour are: site, behaviour, type, and extent of disease.

Site is usually given as an International Classification of Diseases (ICD-8th revision; WHO 1965) code with possible national additions. This classification, however, is constructed from an epidemiological point of view and is not very suitable for the use in an oncology data base. It consists of a mixture of different "axes", and benign tumours are coded with considerably less precision than malignant tumours as regards site.

The idea of describing a tumour in several terms, each describing a specific axis was realised in the Systematized Nomenclature of Pathology (SNOP 1965) which contained four fields: T (topography), M (morphology), E (ethiology), and F (function). Although the entire system has not gained widespread use, the chapter on morphology dealing with neoplasms has been widely adopted (WELLS 1972) but has been found inadequate in certain respects (GRAEPEL et coll. 1975).

The idea of SNOP was carried further in the Manual of Tumour Nomenclature and Coding (MOTNAC) in 1968. Here, topography was based on the ICD code and morphology on the SNOP code. A revision of this classification called ICD-Oncology (ICD-O), specially adopted for oncological needs, is in preparation (TULINIUS 1975). The use of ICD-O is proposed for standardized cancer registries (WHO 1976).

In the current data base the malignant disease is described by the ICD, SNOP, and TNM (UICC 1974) classifications. A more detailed description of the site of the primary tumour, regional lymph nodes involved, and possible distant metastases is provided by a special anatomic code (III). The same code is also used for the description of the target volume in radiotherapy, as well as for critical organs. It may be regarded as an expanded ICD code, specially adopted for radiotherapeutic needs. In the daily routine, this code

has been found adequate and up to now there has not been any need for changes of or additions to the code.

Similar codes have been developed at other departments, e.g. code for the detailed description of ENT-tumours, based on a combination of ICD and TNM, as described by SCHWAB (1975).

TREATMENT DATA

Except for external beam therapy, the different treatment modalities (surgery, other radiotherapy procedures, and chemotherapy) were originally described by a limited number of variables.

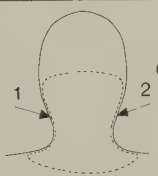
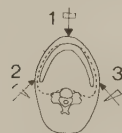
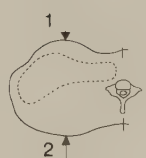
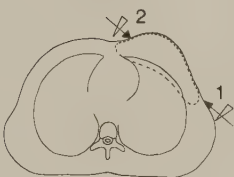
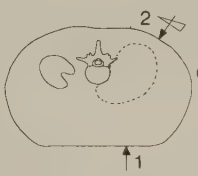
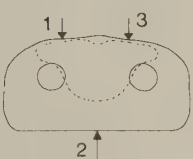
At this Department, since special attention has been devoted to external beam therapy and the development of new treatment techniques, it was found necessary to standardize a number of parameters used in the procedure of delivering external beam therapy, as described in Fig. 1. The definitions comprised the different volumes and absorbed dose values, beam geometry and other factors related to the treatment technique, and the scalings of the treatment units. Since these are described and discussed in detail in III, only a brief summary will be given here.

The target volume is described in terms of the anatomic code as listed and illustrated in III. Two parts of the target volume may be indicated, and for each part, or the volume as a whole, absorbed dose is recorded as maximum target dose, minimum target dose and "representative" dose; for planned treatments the latter dose concept usually equals the minimum tumour dose in the presence of demonstrated tumour.

For the description of beam geometry a special system has been designed: A polar coordinate system is inserted in the transversal cross-section used for treatment planning. The origin of the coordinate system coincides with the midpoint of the midline in the section, and angles are measured from the ventral half of the midline in a counter-clockwise direction, provided that the section is oriented as if seen from the foot-end of the patient. The beam direction is then the incident angle of the central beam, and the entrance point is given as polar coordinates of the point where the central beam enters the patient's contour. The use of this system to describe different treatment techniques is illustrated in Table IV.

With the development of the computerized treatment planning system and the new treatment record form (III), this part of the general data base was expanded and converted to a separate data base, intended to be common to the treatment planning and medical record systems (MÖLLER 1977).

TABLE IV. Descriptions of some standard treatment techniques using the codes for planning sections and the definitions of beam geometry (III).

Planning section and target volume	Code	Field No.	Beam direction		Field coordinates
			in cross-section	in sagittal plane	
 cervical nodes (frontal section)	80	1	90	15	90 - 6.0
		2	270	15	270 - 6.0
 floor of mouth	11	1	0	0	0 - 8.7
		2	136	0	86 - 5.9
		3	224	0	275 - 5.8
 right axillary + supraclavicular nodes	90	1	0	0	50 - 13.2
		2	180	0	129 - 13.6
 left breast	25	1	230	0	276 - 17.5
		2	50	0	339 - 13.2
 right kidney (patient prone)	41	1	0	0	27 - 11.5
		2	135	0	128 - 13.7
 pelvic and inguinal nodes	51	1	0	0	40 - 10.0
		2	180	0	180 - 8.5
		3	0	0	324 - 9.4
 anus or rectum (sagittal section)	87	1	0	270	0 - 0

A similar system for the exhaustive description of intracavitary therapy is intended.

With the rapid widening of the indications for chemotherapy, the originally defined variables are not considered sufficient, and a more comprehensive treatment-specific system, similar to the one described by MESEL et coll. (1976), is being investigated.

FOLLOW-UP DATA

The patients treated at the Department are followed according to the principles described by LINDGREN (1962).

It is desirable, that the patient's condition and the state of the tumour are recorded separately, and that the current status is given in relation to the status initially. The patient's condition may be described by a performance index, and the state of tumour as consecutive TNM classifications as suggested by SPIESSL et coll. (1971) and in the proposed Danish RADATA system (1972).

Entry of the next date of appointment permits revisit planning and the summoning of patients who fail to attend (KAHR & MOSKON 1965).

After the conclusion of personal follow-up, a mailing routine is utilized to obtain annual information on the condition of the patient. This may be performed by the automatic print-out of questionnaires on "Data-mailers", as described on page 27, and also suggested by EDWARDS (1972).

At the autopsy, the state of every tumour diagnosed during lifetime should be assessed, if possible. Any findings of additional malignant disease as well as possible revisions of diagnosis should also be recorded.

If the information on the patient's death is obtained by the mailing routine only, a considerable time lag in the updating of the data base may be experienced. Therefore, it is planned to update the main register file with information on deaths received from the National death record file at regular intervals, in a similar way as the one described by BOLANDER (1975).

DISCUSSION

The definition of the contents of a data base, applicable to patients with any type of malignant disease treated at a department of radiotherapy, has taken place concurrently in a number of countries. In the United Kingdom, a Joint Working Party on Computers in Radiotherapy was initiated by the British Institute of Radiology and the Faculty of Radiologists. In 1969 a first

recommendation of a Minimal List of Variates to be included in a radiotherapy data base was presented (SNELLING & KUNKLER 1970), and a second, revised version was published in 1974. In Denmark, a similar list of variates was defined and nationally accepted (RADATA 1972). Based on the data base, as defined here, discussions took place in the Swedish societies in 1969-70 but a national agreement was not reached. A revised version of this data base, intended for the use in a regional tumour registry, was published in 1973 by ARNESJÖ et coll.

When comparing these different data bases a great concordance is noted. The most obvious deviations are found in the part describing external beam therapy; here the Minimal List defines four separate items depending on the target volume and technique, the RADATA gives the target volume in terms of TNM categories, and in the present system the target volume is described in greater detail by using a special anatomic code.

Systems intended for the collection and processing of general oncology data have also been described by KAHR & MOSKON (1965), ELLIS (1968, 1970), MINET & TIBAUUX (1970), and MINET (1974), GARY-BOBO & ROSSI (1970), OLSON et coll. (1970), BRUCE et coll. (1970), McAINSH et coll. (1971), BRINCKER (1971), EDWARDS (1970, 1972, 1974), NEURATH et coll. (1974), STEBEN et coll. (1976), FOLLAND et coll. (1976), FAYOS (1976), LASZLO et coll. (1976), WHEELER et coll. (1976), among others. In these systems, the number of variables collected are usually less than those proposed in the Minimal List, and more deviations between the different departments may be noted. However, still the items selected are basically the same.

Thus, it is natural that the idea of an international radiotherapy data base has been forwarded (PREWITT 1976). However, before an international agreement is reached on the definitions of important variables such as target volume and absorbed dose (ICRU Report, in preparation) and the problem of accuracy in reported dose values is solved, conclusions drawn from calculations based on such a pooled material may be quite misleading (LANE et coll. 1976).

The data base, as originally defined for a department mainly devoted to radiotherapy, must be flexible enough to accomodate additional data which may be desired when the departments of radiotherapy are developed into departments of oncology and oncological centres. For this purpose, the addition of disease-specific files or data bases devoted to e.g. clinical trials, as suggested by ARNESJÖ et coll. (1973) and LASZLO et coll. (1976), may be suitable.

Quite a number of disease-specific data bases have been described: brain

tumours (KRICHEFF et coll. 1970), head and neck tumours (SCHWAB 1975), tumours of larynx-hypopharynx (MacKINNON 1970), breast (GROS 1970, BARCLAY 1975), gynaecologic neoplasms (SNELLING 1970), urologic neoplasms (DEELEY 1972, WICKHAM et coll. 1975), Hodgkin's disease (PARZIALE & KAPLAN 1973) to mention a few. In some systems a number of site- or disease-specific proformas are used to describe the disease, and one or more general proforma(s) for the recording of treatment and follow-up: PHILLIPS (1972), LASZLO et coll. (1976).

The treatment documentation system developed at this Department constitutes a treatment-specific data base. In addition to the evaluation of given therapy, the system was intended to give an exhaustive description of treatment fulfilling the needs of the physician in his daily work; to provide the radiographers with daily operational data; to enable a description of the work at the different treatment units not only in number of patients and treatments but also in complexity of setting-up, which may yield information leading to improvements in technique, working routines, and environment and increase in staff; and to provide the physicist with information on the use of different types of shielding, collimators, and other physical parameters. This will be possible by the merging of data from planning, scheduling, simulation, setting-up and control measures (III).

The same approach to an integration of medical and physical parameters have been discussed by e.g. DUTREIX et coll. (1972).

FUNCTIONS OF DATA RETRIEVAL

Three physically separate files are concurrently used for the retrieval of information: the main patient record file, the fixed format "statistical" file, and the register file of the treatment documentation system.

MAIN PATIENT REGISTER FILE

In this file all information collected at present according to the definitions of the general data base, except for details on external beam therapy, is stored. Due to its size, about 25 000 records, with an average length of 700 characters, only a few retrieval functions operate directly on this file.

Direct access routine

This routine, as described in II, was created chiefly for checking, correcting and editing purposes. The main patient record file was transferred to a disk file, on which a VDU was operating. Any record could be searched by its CRN or the initial 14 characters of the name. Every single variable could be changed or deleted, and additions could be made. The corrected data was temporarily stored in a transaction file, which updated the main register file at regular intervals.

The access time of any record was usually less than two seconds. For each correction of a single variable three or four separate commands were necessary.

Patient listings

A separate file was created at each updating to enable listings of the complete patient material (II). Different types of lists may be chosen: patient identity lists (CRN, name, title, address and parish) sorted on CRN or family name, and diagnosis lists (CRN, name, diagnosis No., date of admission, ICD code, clinical diagnosis text, SNOP code and microscopic diagnosis text) sorted on CRN, year of admission, ICD code or SNOP code.

The rationale behind these listings was to provide the secretaries and staff at reception with an up-to-date patient register for reference purposes.

The diagnosis lists were also used for the manual checking of the ICD and SNOP codes.

Mailing follow-up routine

A special routine, which enables the automatic print-out of follow-up questionnaires on "Data-mailers" was designed. Such letters may be addressed to:

- a. The patient. In this case the questionnaire contains a number of questions relevant to the disease, for which the patient was treated.
- b. The hospital or physician currently caring for the patient, in which case a standard questionnaire is printed, containing questions on whether the patient has been attending during the time interval, his condition and, if deceased, the date and cause of death. If the patient is not recurrence-free or is deceased, a complete copy of the pertinent parts of the patient's record is asked for.
- c. The registrar's office in the parish of the patient. This routine is used when no answer is obtained from the patient and for elderly patients. The questionnaire is standardized and contains questions whether the patient is alive or dead, his present name and address, or date and cause of death.

The answers to the questionnaires updates the main record file.

"STATISTICAL" FILE

A special file in fixed format is generated at updating (II). This file consists of a strictly specified set of variables extracted from the main register, with a record length of 132 characters, one record per diagnosis. The contents of the file are described in Table V, in which also the various search keys are listed. At present, the file is stored on disk at the local computer, which enables immediate access to the file via a VDU or typewriter terminal.

Search program

The key words listed in Table V are used for search in the "statistical" file. The maximum number of keys at a single search session is ten. Any key may be given as logical "NOT". The search may be terminated when the actual number of records satisfying the search profile have been obtained, but it is also possible to proceed with other program modules, as indicated in Fig. 5. The most simple form of output is a listing of the contents of the temporary file, resulting from the search operation. The list has the same format as the

TABLE V. Structure of the records of the fixed-format file.

Position	Variable	Search key word	Type
1 - 10	Patient's CRN	ÅLDER	4
12 - 28	Name	-	
30 - 35	Parish	FÖRSAM	4
37 - 38	Diagnosis No.	DXNR	2
40 - 45	Date of referral	INREM	4
47 - 51	Referring hospital	LAS	4
53 - 55	Referring department	KLINIK	4
57 - 62	Date of admission	INKOM	4
64 - 68	ICD8-code	WHO	4
71 - 74	Clinical stage	STAD	3
76 - 78	T-category		
79 - 81	N-category	TNM	3
82 - 84	M-category		
86 - 89	SNOP-code	SNOP	4
91	Confirmation of diagnosis	HISTYP	2
93	Not accepted for treatment	REFUS	1
95	Diagn. prior to adm.	TIDIAG	1
97	Radiotherapy-external beam	EXT	1
101	Radiotherapy-intracav. etc	INT	1
105	Radiotherapy-metabolic	MET	1
109	Chemotherapy	KEM	1
113	Follow-up status	FU	2
115 - 120	Date of death	MORS	4
122	Cause of death	MORSAK	2
123	State of primary at death	TUMKV	2
124	State of metastases "	HETAST	2
126	Autopsy performed	OBD	1
129 - 131	Survival, weeks	SURVIV	4

Type 1 indicates a Boolean variable, 2 a value,
3 a valustring, and 4 an interval.

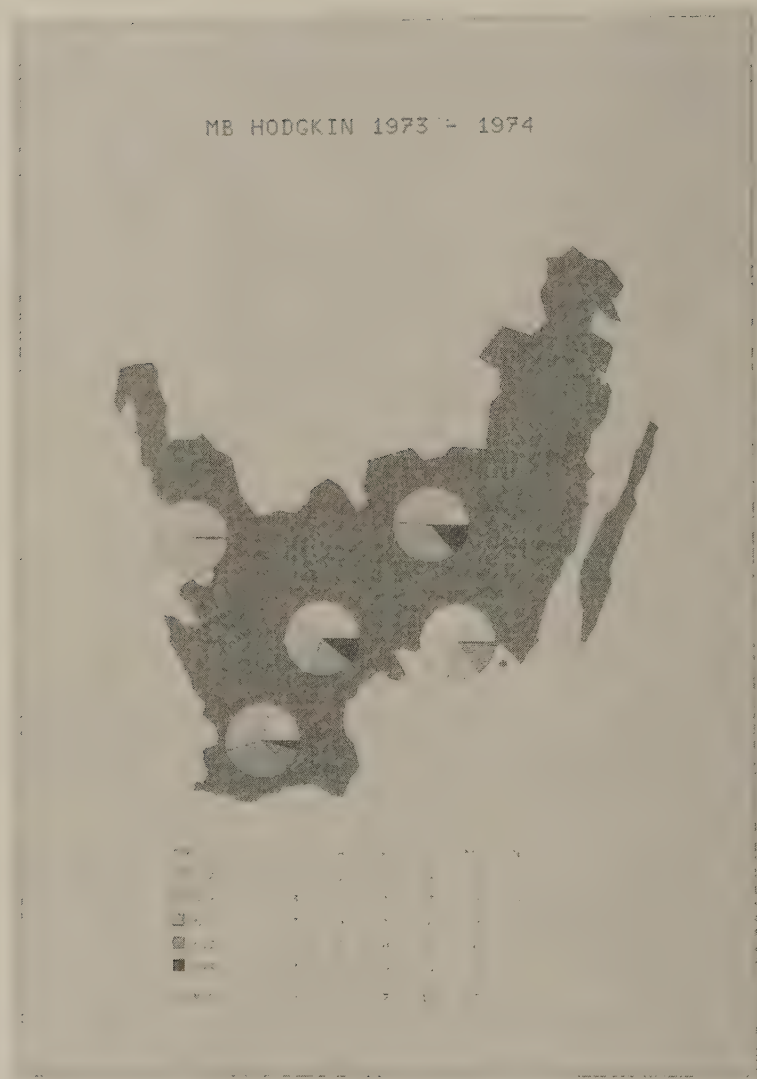


Fig. 6 Colour presentation produced by a jet ink plotter. The map shows the area of the southern part of Sweden, served by the Department of Oncology in Lund. For each county, the distribution of the various subtypes of Hodgkin's lymphoma is given by the sectorial diagrams, and in the table underneath.

Any one of these tables may be used separately, e.g. age distribution, on any patient material satisfying other search profiles. Some of these programs may also generate a data file suitable as input to the colour display system described below.

Programs for the presentation of annual statistics on the different treatment modalities were also developed.



Fig. 7 Colour presentation of incidence of mammary carcinoma in the County of Malmöhus, given for each community. Since the figures are based on the number of cases referred to the Department of Oncology in Lund, they are not representative of the true incidence in the different areas.

Incidence

Annual census data for each community in the area served by the Department of Oncology are stored in a separate file (Fig. 5). An interactive program, utilizing the statistical file as well, enables listings of the incidence rates for any type of tumour in any age interval for any geographical area. These programs have been developed merely for use in a future regional tumour

registry, since the incidence rates obtained today only reflects the incidence of malignancies as met in the Department of Oncology and not the true incidence in the region as a whole (Fig. 7).

Colour display

A special device has been developed at the Lund Institute of Technology producing multicoloured pictures by means of electrically controlled ink jets. The picture is produced in 30 seconds and has a resolution of 0.2 mm. Areas may be plotted in arbitrary colour and shading pattern, and a standard program package permits easy production of histograms, sectorial diagrams, or 3D graphs (JERN 1976). Such presentations may be superimposed on e.g. maps.

Some of the program modules described have this facility of display (Figs. 6 and 7).

TREATMENT DOCUMENTATION SYSTEM

The output from this register has been described in detail in I. The main feature is the daily print-out of treatment reports on all patients, whose external beam therapy has just been concluded. These reports are included in the patient's case notes. A more comprehensive report may be obtained when needed (for example if the original treatment record form is lost). Included are also programs for evaluating the quality and completeness of data collection, and for annual statistics as regards workload on the different treatment units.

In the batch processing systems (see page 41) the selection of a group of patients with a specified treatment technique necessitated search in separate register files. With the conversion of the total system to the local computer data base this difficulty will be overcome.

DISCUSSION

With the initially defined goals of the system (page 9) as a starting point, the functions of data retrieval were designed.

The intention of easy access to an up-to-date patient register for reference purpose was initially met by the patient listings. These were, however, only produced at irregular intervals due to their size (about 1 000 pages) and the cost of producing them. To get a more manageable form of presentation,

microfiche was considered but would be even more expensive. With the implementation of the direct access routine a satisfactory routine was established, replacing the patient listings.

With the change of central computer, however, the operation of the routine was restricted, and finally had to be discontinued. The local interactive computer system will enable the optimal use of direct access.

The search programs were developed to make selection and listing of various groups of patients possible. This routine has entirely replaced the filing cards previously utilized by the secretaries for the retrieval of patients' records. The number of searching procedures is increasing, and several are performed daily. The format of the listing has in most cases been found adequate, giving the information desired. Similar search programs have been described by e.g. FREY (1969) and STRÜTER & HALAMA (1974).

The use of a number of program modules, operating on the same input file, is a flexible arrangement permitting combination of different modules at a single search run. New program modules can easily be added.

The annual statistical report has a definite value in giving a standardized description of the patients referred to the Department during one particular year. The geographical distribution of the patients and the number of patients referred from different hospitals may give information on the current treatment policies in the region. The time lag between referral and admission is given. Changes in the pattern of diagnosis and age distribution may be demonstrated. The report is also part of the checking procedures.

The colour display has been extensively used in technical sciences such as architecture, town-planning, geography, and geology but its application in medicine has so far been limited. Besides the use illustrated here, it has been utilized for the presentation of scintigraphic images after picture processing (GUSTAFSSON 1973) and for chromosome analysis (BODOR et coll. 1974). A colour display, combined with graphs of different types, can present information else given in tabular form, thus enhancing readability and facilitating interpretation.

According to recommendations by UICC (1974) survival should be calculated from the date of start of treatment or date of decision not to treat. In this system, however, survival is calculated from the date of admission as recommended by WHO (1976), since this date is recorded for all patients, and the difference was considered negligible. Patients treated prior to admission should be analysed separately. The curves presented are similar to those described by FAYOS (1976).

For the analysis of cure rate, recurrence rate, and side-effects of treatment, as discussed by e.g. HUTCHISON (1973) and FONDARAI (1974), the

continuous recording of follow-up data is necessary. Since the system does not have the capacity of recording such data at present, no specific programs for the calculation of these rates have been developed. However, with the conversion of the system to the local computer the possibilities for the collection of follow-up data are currently investigated.

VALIDITY CHECKS

The quality of information in medical records has been discussed by WAGNER (1970), FEINSTEIN (1970), BLEEHEEN et coll. (1970), among others, and it is generally agreed that such information is much less reliable than is commonly believed. The introduction of computers does not in itself raise the quality of information, but the need for standardized descriptions and adequate definitions of data to be recorded have a positive effect on the quality of data, and a educational effect on the staff. The computer can be programmed to perform a great number of controls and checks of the recorded data, which can be repeated endlessly without loss of precision, in contrast to manual procedures.

Two types of errors may be distinguished: factual errors and formal errors (WAGNER 1970). Factual errors are false statements or incorrect entries, which can only be detected by going back to the source material. Computer checking of factual errors is difficult and often impossible, whereas formal errors, such as the incompleteness of records, the excess of limits, the occurrence of impossible codes, and inconsistencies of all kinds may easily be checked by appropriate computer programs. Such checks should be performed preferably already in the process of data collection.

COMPUTER CHECKING PROCEDURES

For the current system, the type of data collection technique used permits immediate checking of the recorded data. However, at the moment only the civic registration number is checked. This was considered most important, since record linking is based on the CRN, and if erroneously recorded, information pertaining to different individuals may be mixed. The check is performed by calculating a check digit and comparing the result with the recorded check-digit (page 19). Since this method is insensitive to inversion in the order of pair of digits, it is, however, not infallible (BOLANDER 1975).

Another check is the automatic shifting of colour on the typewriter ribbon, which indicates that the recorded data is accepted by the system, and also serves the function of indicating that data recording for computer processing

has taken place (II, IV).

At updating the main register, all new transactions are thoroughly checked as regards date, chronology, and codes and all errors are listed (II). Similar routines have been developed for the treatment documentation system (I).

MANUAL CHECKING PROCEDURES

Manual checking, where the complete contents of the data base for a specific patient are checked against the medical record, is not performed regularly. However, since the system is often used to produce lists of various groups of patients for manual retrieval of case note data, the contents of the lists are usually compared to the information in the case notes, and if any errors are found they are corrected.

The patient listings, containing the full text of clinical and microscopic diagnoses (page 26) are very suitable for the checking of the validity of ICD and SNOP codes.

In the programs for presenting annual statistics, each table has a specific set of search parameters. Records not fulfilling the search criteria of a specific table due to e.g. format errors or omissions, could be listed. This method, for example, revealed the unexpected referral of a male patient from the Department of Gynaecology; it was indeed a new-born boy with a hemangioma.

VALIDATION STUDY OF THE PRESENT SYSTEM

To get an estimation of the overall accuracy of data recording and quality of data contained in the main register an external validation study (MARTIN et coll. 1976) was performed. A sample of records on patients with malignant disease was obtained from the main computer record file by means of a random number generator. As the basis for selection, 1974 as the first year of admission was chosen. As a result, 240 patient records were retrieved and listed, and these computer-produced records were carefully checked against the case notes of the patients. The case notes of three patients, however, could not be found despite considerable effort, which is per se an error in the manual record system. Thus the comparison was based on 237 records. The results of this study are presented in Table VI. The sample had the same profile as regards distribution of age, sex, and diagnosis as the total 1974 patient population as listed in the internal annual statistical report.

TABLE VI. Results of external validation study of 237 patient records selected at random.

	Total No. of subrecords	Error rate %	No. of conversion errors		No. of input errors		
			Erroneous subrecords	Missing subrecords	Errors in primary data	Spelling and typing errors	Not recorded
Civic registration number	237	-					
Name	237	2.1		3		2	
Address	237	3.4		2	1	5	
Parish	237	1.7		2		1	1
Referral, date and department	244	0.4				1	
Admission, date	261	-					
Diagnosis, text	261	4.2	1		4	6	
ICD code	261	1.9	1		3	1	
PAD/CD, text	255	3.1	1			6	1
SNOP code	255	3.1		2	1	3	2
Stage	53	5.7		2	1		
Radiotherapy, external	408	9.8	1	15	3	5	16
"-", intracav.	68	11.8		2	1		5
"-", metabolic	5			1		1	1
Chemotherapy	68	14.7		4		1	5
State of follow-up	41	14.6		2			4
Date and cause of death	82	6.1		4			1
Autopsy performed	82	4.9		2			2
	3292	(3.9)	4 (0.1)	41 (1.2)	14 (0.4)	32 (1.0)	38 (1.2)

A total number of 3292 items or subrecords were checked. Out of these, 3 163 were correctly retrieved from the computer file, whereas 50 showed errors of various kinds and 79 were missing. Errors at the input stage were found in 84 subrecords and could be attributed to wrong primary data in 14, typing errors in 32, and omissions in 38 cases. Conversion errors, i.e. errors in the chain of recording, conversion, transaction generation, and updating were found in 45 cases, with 4 defective and 41 missing subrecords.

The total error rate was 3.9 %. The error rates for the different types of subrecords are given in Table VI.

The typing errors were in general negligible, consisting chiefly of spelling mistakes or mistypings in the text strings of address, and clinical and microscopic diagnoses. In five instances only these errors were of real importance; one giving a false code for parish, and the others a wrong ICD or SNOP code. With an average subrecord length of 20 characters, typing errors amounted to 1 in 2 000 characters.

Errors in the primary data consisted mainly of a wrong clinical diagnosis text and ICD code given by the physician, and in three cases a wrong series No. for external beam therapy.

The highest omission error rates were observed for the recording of given therapy. The reason for this is probably forgetfulness of the personnel at the treatment units since this entry is based on a separate notification to the ward (or out-patient department) that radiotherapy has commenced. During the study,

all treatment records and computer-produced treatment reports included in the case notes were checked, and in no case any defects or omissions were noted. Thus, the treatment documentation system compensated entirely the omissions noted for recording that treatment modality.

For the majority of cases where recording of chemotherapy was missing, the treatment had been started at another department prior to admission or during follow-up. Conversion errors were noted for practically all types of subrecords. However, two thirds of these errors occurred after the change of central computer, which introduced some program errors.

DISCUSSION

A multitude of factors may influence the quality of data contained in and retrieved from a computerized information system: type of data collection system, instructions for use, attitude of staff, design and reliability of computer software and hardware.

Inadequate source material, usually the information in the patients' case notes, may have a deleterious effect on the final quality of data. In the analysis of data retrieved from case notes of patients with laryngeal and hypopharyngeal carcinoma undertaken by BLEEHEN (1970), errors ranging up to 40 % were noted for information on symptoms, but also the description of site and extent of primary lesion as well as assessment of result of treatment showed a high error rate. It was concluded, that major contributions to these errors were to be found in the ambiguity of records, and to changes in the international nomenclatures. As a consequence of these findings, a set of proformas was designed for the actual diagnoses, replacing the ordinary patient record, and at the same time being the method of data collection for computer processing (MacKINNON 1970).

In a survey of information retrieved from the computer of a large abstracting service, HENDRICKSON & MYERS (1973) found "an extremely high rate of incorrect entries". It was concluded that the data was unsuitable for quality care studies, either due to a very high error rate or recording bias. These defects could be attributed to inadequate training of the coders, unintelligible instructions and lack of checking procedures.

Such observations stress the importance of organizing the medical record, the use of agreed vocabulary and standardized definitions on terms, and the collection of information as near to the source as possible, avoiding unnecessary data transfer and keeping coding to a minimum.

The error rates for different input systems are not easily compared.

Ordinary key-punching, with verification, usually has an error rate of 1 in 1 000 characters (BLEEHEN 1970). In the collection of free or narrative data from pathology reports, OKUBO et coll. (1975) found punching errors in 6 - 7 % of the entries. If screening of the documents before punching was not made, spelling error rates of 20 - 25 % and format error rates of 8 - 9 % were experienced. KOREIN (1970) reported 6 % typographic errors in the recording of date in 100 test cases, and in an extensive internal validation (formal and semantic) study MARTIN (1976) found that 6 % of all errors arose from mistakes made at the keypunch stage. The rate of spelling and punching errors found in the present system thus compares favourably with these figures, and may probably be attributed to the competence of the secretarial staff.

The various input devices used may introduce errors, but card readers and optical character readers are usually technically reliable. Some difficulties have been experienced with the tape cassettes involved in the off-line data collection system. After repeated use, they often were rejected by the converter, and the recorded information had to be copied on a new cassette. The conversion error rate in Table VI also includes programming errors in decoding the information, but is still rather small (1.3 %), as compared to a 4 % format rejection rate reported by KOREIN (1970), using the same type of equipment. With the on-line data collection technique, errors introduced at conversion of the cassette-tapes are eliminated.

A source of error, perhaps not usually appreciated, is the international coding systems. The different classification axes used in ICD may lead to ambiguity or loss of precision, and not all tumours can be coded by SNOP. The asynchronous revisions of ICD and TNM may introduce errors and the different vintages of the classification systems are not always comparable. In a study including Scottish morbidity data, GREENWOOD (1972) found a distressingly high percentage of errors in manual disease coding, 20 %, as compared to errors in computer coding of the same diagnostic phrases, 2.4 %. Although the present study did reveal only three erroneous ICD codes, it would be worth considering the use of such programs in the checking procedures.

The checking procedures applied to the treatment documentation system are rather elaborate as regards formal errors, and result in treatment reports of high reliability as shown in the study. However, factual errors may present problems now and then, especially in the correct description of the target volume. Education of staff and the use of agreed anatomic terms have greatly reduced the frequency of that particular error.

The accuracy of the absorbed dose values entered on the treatment record form is guaranteed by the use of in vivo dose measurements and other control

measures, as described in III. For the present system, it may be stated, that the accuracy of absorbed dose values in external beam therapy is in the order of $\pm 5\%$, and in some special cases even better.

Attention should also be paid to the time lag in the recording of data and updating of the register, as discussed by BRINCKER (1977). If the data base is updated only once annually, this will affect the validity of e.g. recurrence rates obtained from the system.

The following measures are proposed for reducing the number of errors:

- a. Additional checks performed at the recording stage;
- b. Daily updating of main register with printing of list of errors. The patients' case notes will then still be current and may be easily collected for completing and correction purposes;
- c. Prompt and efficient correction routines, preferably on a VDU in direct access;
- d. Frequent collation of data from separate files, e.g. main record file and treatment documentation file;
- e. The design and implementation of special systems for different treatment modalities;
- f. Frequent updating of main register file with information on death obtained from the National death record file;
- g. Processing of free or narrative text in special strings.

A number of these additional checking procedures is in fact planned and will be implemented on the local interactive system.

DATA PROCESSING

Since the start of the project in 1967, the hardware and software employed have changed considerably. This has been due to the type of computer available and its operating system, as well as to local administrative and economic considerations.

BATCH SYSTEMS

Initially, data collected on the proformas previously described was punched on paper tape, and the tapes were transported to the computer, a SAAB D21, located in a neighbouring town. The processing was performed with J5, a system for processing medical information developed by HALL et coll. (1967). A satisfactory level of functioning, however, could not be attained with this system, although some record print-outs were obtained.

The structured radiotherapy record was implemented in clinical routine in 1970, and data collection on the MT 72 off-line system (II) was started in 1971. The cassette-tapes were converted to standard magnetic tape at a commercial service bureau, but processed at the Univac 1108 computer at the Lund University Computer Centre. This large central computer, initially run under the Exec 2 but later under the Exec 8 operating system, was well equipped with mass storage and varying types of peripherals. The medical information was processed by an information system developed by SUNDSTRÖM (1970 - 1975). Variable field length format, organized sequentially, was employed (KOREIN et coll. 1963). Each subrecord consisted of a header of 18 characters describing the length and type of the subrecord, diagnosis No., and an internal code, and a variable value-field.

In 1971 a remote job entry terminal (Univac 9200/II), equipped with paper tape reader, card reader, line printer and incremental plotter was installed at the Department. This terminal greatly facilitated the development of programs for information retrieval and some of the programs operating on the fixed-format file were developed. Concurrently, systems for the processing of scintigraphic images (MÖLLER et coll. 1972) and radiation treatment planning (III) were designed.

The system for documentation of external beam therapy (I), was implemented in 1972 and was run in a daily routine with print-outs of treatment reports.

For administrative reasons the central computer was changed in 1975 to a Univac 1110, located at the computer centre of Malmöhus County Council, and most of the programs had to be transferred. This computer, however, was oriented towards administrative data processing, and the configuration differed from the one previously employed.

DIRECT ACCESS ROUTINE

With the installation of a VDU (Uniscope 100) in 1974, it became possible to perform corrections directly in the main record register. For this purpose, the sequential files were converted to a format, suitable for direct access. The transactions generated via the demand terminal were stored in a separate file, updating the main register at regular intervals (II). Interactive programs operating on the fixed format file were also developed, and the results were presented on the printer or plotter of the remote batch terminal.

The use of this facility had to be discontinued in 1976 due to the high running costs.

LOCAL INTERACTIVE SYSTEM

In the long run, remote batch processing did not meet with the demands of the users: the turn-around time was too long for image processing and treatment planning, the computer was only accessible during working hours, and the discontinuation of the direct access routine deprived the record system of its main facility for correction. In addition, the running costs did not permit any further development of the three systems, and thus it was not possible to obtain an interactive type of processing.

Therefore, in 1975 the remote batch terminal was replaced by a local computer, PDP 11/45, still linked to the Univac 1110. A stepwise expansion of the local installation was scheduled; during the first year the nuclear medicine system was converted to local processing, and during 1976 - 77 the dose planning system and the medical record system are being converted to local service.

The current configuration of the local computer system is shown in Fig. 8. The RSX-11M operating system is used. The programs are written in Fortran

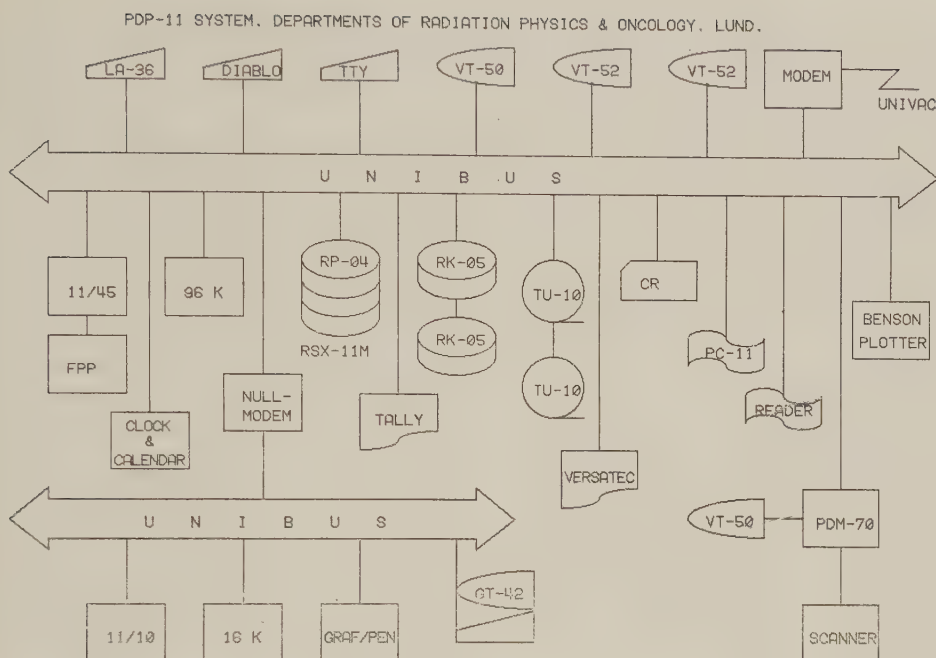


Fig. 8 Current configuration of local computer installation. Main unit (top) is a PDP 11/45 with 96 kwords (16 bits) parity-memory. A second unit (bottom) is a PDP 11/10 equipped with a graphics display, used for interactive treatment planning.

IV and, for some special applications such as the data collection system, in Macro-assembler. In the third phase of expansion, the configuration will be completed with a PDP 11/34 running under the MUMPS-11 data base operating system. This is necessitated by RSX and MUMPS being different operating systems, but capable of communicating and exchanging data.

A comparison of the different computer systems is given in Table VII.

DISCUSSION

The various types of computers, which may be utilized in radiation therapy treatment planning, have been discussed in III. The conclusion was that the medium-sized multipurpose computer alone or linked to a large central computer, offered the best flexibility and greatest computational potentials. During the last few years, an increasing number of such installations have been reported (III).

For the processing of medical radiotherapy data, most existing systems are run on a large computer in batch mode. In the survey of computers and

TABLE VII. Comparison of two computer systems employed.

	Remote batch system		Local computer system
Computer	Univac 1108/1110		PDP-11 system
Operating system	Exec 8	RSX-11M	MUMPS-11
Programming language	Fortran (compiling)	Fortran (compiling)	Mumps (interpretative)
Turn-around time for program modification	1 h	2 min	None
Main record file storage medium	Magtape	Not intended	Disk
Organization of files	Separate files on separate magtapes	-	All "files" within the same data base
File structure	Sequentially organized	-	Hierarchical data base
Turn-around time for retrieval from fixed format file	30 min - 2 h	15 - 20 min	Not tested
Accessibility of main file by direct access routine	1.5 h 3 times weekly	-	No restrictions
Retrieval of individual records	By programming	-	Directly
Costs	Lease of terminal: Skr 55 000/year, computer time: Skr 120 000/year	-	Leasing: Skr 90 000/year, service: Skr 30 000/year
Personnel	1/3 programmer 1/3 operator 1 secretary		1/3 programmer 1/3 operator 1 secretary

computer programs in radiotherapy reported by DAHLIN (1972) at the Fourth International Conference in Uppsala, a total of 171 different programs were listed (external beam treatment planning 80, intracavitary and interstitial dose calculations 42, beam data collection 16, treatment automation 11, administrative 22) for 96 computer installations. Of the 22 medical data systems, 7 were run on minicomputers. With the development of MUMPS (Massachusetts General Hospital Utility Multi-Programming System; BOWIE & BARNETT 1976), a powerful tool to handle medical information on a minicomputer became available. Examples of multipurpose radiotherapy departmental computers using MUMPS for record keeping purposes and tumour registry services are given by e.g. NEURATH (1974), STEBEN et coll. (1976), FOLLAND et coll. (1976). A comprehensive list of MUMPS applications may be found in the MUG-3 report (1976).

The lack of success of the initial batch system is more to be attributed to the mode of data collection and input (proformas - paper tape) and the geographical distance between user and programmers, than to the system itself. The J5 system is still used in some special applications at the Karolinska Hospital in Stockholm (MELLNER et coll. 1974).

The data processing system used since 1971, developed at the Lund Univer-

sity Computer Centre, has proven very efficient and reliable. However, the structure of the system is rather complex, and the programs for decoding the structured strings and generation of transactions may be quite large. This makes major modifications of and additions to the data base a matter of rather extensive programming. The implementation of the system on another computer was quite an accomplishment, but errors in the programs were introduced as shown in the external validation study.

Since a medical information system must be dynamic to keep pace with the varying demands necessitated by the development of new and better diagnostic procedures and treatment techniques, attention was focussed on MUMPS as the data base operating system on the local computer. The MUMPS data base is organized in a hierarchical manner, and it is easy to build a new tree or "global" for the inclusion of a new field of interest.

A MUMPS data base structure applied to the present data base is suggested in the Appendix (Table IX). Preliminary studies with the MUMPS system show great flexibility, easy access, and properties suitable for editing and correcting.

GENERAL SUMMARY AND CONCLUSIONS

The unique characteristics of a department of Oncology give rise to certain problems that can only be solved by the use of computer. One of the main problems is the continuous evaluation of given radioterapy, based on long-term follow-up. To achieve the goals defined in the Introduction, development of new techniques and improvement of already existing ones were made.

In the developmental work, three data collection systems were consecutively introduced into the daily routine. The first one (1968) was based on proformas, intended to be completed by the medical staff. The data were punched on paper tape for further processing by computer. This method of data collection did not gain acceptance, neither from the secretaries nor from the physicians, and was therefore abandoned. A second data collection technique was developed (1971) and taken into routine use (II). A prerequisite for that technique was a structured medical record, and such was therefore designed (IV). Data were collected on magnetic tape cassettes when typing the record. This system was used routinely for five years. Although it proved reliable and was accepted by the staff, it did not meet the requirements of flexibility, and was quite expensive.

In order to overcome these problems the routine was replaced by a third data collection system (1976), based on a minicomputer (IV). The experiences so far have shown properties superior to the others as analysed in Table II. This improved technique was also applied to data collection from the treatment record form (Table III).

The contents of a medical data base common to patients with any form of malignant disease were defined. Compared to other oncological data bases reported, the contents were found to be basically the same.

At this Department interest has been focussed on the development of treatment techniques for external beam therapy. Concurrently with the definition of the data base, a number of treatment parameters were defined (III). With the design of a treatment planning system in 1971-72, a special documentation system for external beam therapy was developed (I, III). The data recorded here form a specific data base, common to the treatment planning and the medical record systems.

A number of data retrieval functions have been developed and described in

this summary.

The validity of the data contained within the data base has been discussed. A special validation study of patient record data was performed (Table VI). A total error rate of 3.9 % was found. The most contributing factors were: omissions (1.2 %), conversion errors (1.3 %), and typing errors (1.0 %). The significance of the different types of error has been discussed briefly, and measures to reduce the error rates have been suggested.

A comparison of the different computer systems employed favours the use of a departmental minicomputer. (Table VII).

Based on the experiences and results of analyses presented here, the following conclusions may be drawn:

- ☐ Structuration of the radiotherapy record is quite feasible, with positive acceptance of the medical staff and the secretaries;
- ☐ The interactive, on-line data collection system is reliable, fast, flexible, and user-oriented;
- ☐ The collection of data relevant to external beam therapy is smoothly performed directly from the record form. An interactive typewriter terminal or VDU input is faster and superior to key-punching as regards operator's errors;
- ☐ The proposed system for description of external beam therapy has proven adequate for any type of technique used at this Department. The anatomic code has been found satisfactory;
- ☐ The contents of the data base are in accordance with other systems proposed. It may easily be converted to a MUMPS data base structure, if desired;
- ☐ The retrieval functions developed so far have been found valuable to the medical staff, the secretarial staff, and other personnel involved;
- ☐ Compared to a large central computer, the departmental minicomputer system was found to have a greater flexibility and ease of access, and to be more economic.

ACKNOWLEDGEMENTS

My most sincere thanks are due to:

Professor Martin Lindgren and Professor Kurt Lidén for initiating the study and for their support and encouragement;

Carl-Håkan Håkansson for valuable help and advice;

Thomas Gustafsson, John-Erik Johnsson, Torsten Landberg, Ulla-Brita Nordberg, Arne Sundström, and Gudrun Svahn-Tapper for their competent and pleasant collaboration;

The staff at the departmental computer for programming assistance;

The secretarial staff for their obliging assistance;

Eva Henriksson for drawing the figures and for skilful secretarial assistance;

My thanks are also due to the staffs of the Departments of Oncology and Radiation Physics.

This work was supported by grants from SPRI (Project No. 210), John och Augusta Perssons stiftelse för medicinsk vetenskaplig forskning, and departmental funds.

APPENDIX

In the Appendix, the proformas designed for the collection of information on patients with any type of malignant disease, are given. The corresponding part of the medical data base is listed for each proforma.

As described, the proforma for documentation of external beam therapy (Fig. 11) was replaced by a special treatment record form, which is described in detail in I and III.

The data base has been subject to revision. These changes are, however, small and do not affect the data base in principle. Even the extended data collection for the documentation of external beam therapy does not alter any of the variables defined initially.

For a revised version of the data base, intended for use with the MUMPS data base operating system, three separate globals have been outlined: identity, diagnosis and mors (^ID), external beam therapy (^EXT), and follow-up (^FU).

It is planned to add separate globals for the description of intracavitary and interstitial therapy (^INT) as well as chemotherapy (^KEM).

TABLE VIII. Contents of the general data base.

PERSONAL DATA

Civic registration number (CRN)	yymmdd-nnnc
Name	Text
Address	Text
Marital state	z
Occupation	Text
Parish	Text + code

ADMINISTRATIVE DATA

Date of referral	Date
Referring dept./physician	Text + code
Admission to ward	Date
Discharge from ward	Date
Visit to out-patient clinic	Date
Not accepted for treatment	*

DIAGNOSTIC DATA

Clinical diagnosis	Text
Microscopic diagnosis	Text
Method of confirmation:	Clinical only/ X-ray diagnosis
Biopsy, histopathology	Specimen-No. + date
"-", cytology	Specimen-No. + date
Diagnosis	ICD-8 code
	SNOP code
Site of primary	Anatomic code (III)
Regional nodes involved	No/yes + anatomic code (III)
Metastases	No/yes + anatomic code (III)
Clinical stage	0/I/II/III/IV, a/b/c
TNM cathegory	Tz Nz Mz

MEDICAL HISTORY

Heredity:	
Twin	No/yes/not known
Ca among relatives	No/yes/not known
Previous malignant disease	No/yes + year, diagnosis, dept.
Previous radiotherapy	No/yes + year, target, dept.
Present illness:	
Duration of symptoms	zz months
Symptoms	Code

TABLE VIII. Cont'd.

SURGERY

Operation No.	z
Date of operation	Date
Type of operation	Text + code
Department of surgery	Text
Microscopic examination	Specimen No. + date
Microscopic diagnosis	SNOP code
Radicality	No tumour demonstrated/tumour, radical excision/tumour, radicality uncertain/tumour, not radically removed/tumour, radicality not possible to evaluate.

Formulärsnr B2 81 Anmald till behandling år mån dag 82 Behandling påbörjad 83 avslutad Planering: Maj 84 Juni 85 Beh. serie nr: 					
RADIOLOGISKA KLINIKEN I LUND					
EXTERN STRÅLBEHANDLING (DISTANSBEHANDLING)					
11 Fält nr 20 Target 21 + Plan genom 22 Strålkälla 23 + Teknik 24 FHA (FAA) cm 25 Fältstorlek 26 x cm 27 x cm 28 x cm 29 x cm Fältkoordinater, °-r 31 - 32 Strålriktning, ° 33 Filter 300 31 301 Antal beh. 35 Fältdos i max, rad 38 	12 Fält nr 20 Target 21 + Plan genom 22 Strålkälla 23 + Teknik 24 FHA (FAA) cm 25 Fältstorlek 26 x cm 27 x cm 28 x cm 29 x cm Fältkoordinater, °-r 31 - 32 Strålriktning, ° 33 Filter 300 31 301 Antal beh. 35 Fältdos i max, rad 38 	13 Fält nr 20 Target 21 + Plan genom 22 Strålkälla 23 + Teknik 24 FHA (FAA) cm 25 Fältstorlek 26 x cm 27 x cm 28 x cm 29 x cm Fältkoordinater, °-r 31 - 32 Strålriktning, ° 33 Filter 300 31 301 Antal beh. 35 Fältdos i max, rad 38 	14..... Fält nr 20 Target 21 + Plan genom 22 Strålkälla 23 + Teknik 24 FHA (FAA) cm 25 Fältstorlek 26 x cm 27 x cm 28 x cm 29 x cm Fältkoordinater, °-r 31 - 32 Strålriktning, ° 33 Filter 300 31 301 Antal beh. 35 Fältdos i max, rad 38 		
Σ: TD max, rad 39 TD min, rad 40 TD medelrad 41 plan. 42 mätt 43 					
<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; vertical-align: top;"> Kritiskt organ I 44 Dos, rad Plan 45 Mitt 46 " " II 47 " 48 49 " " III 50 " 51 52 </td> <td style="width: 50%; vertical-align: top; padding-left: 20px;"> Antal dygn 53 Antal fraktioner 54 </td> </tr> </table>				Kritiskt organ I 44 Dos, rad Plan 45 Mitt 46 " " II 47 " 48 49 " " III 50 " 51 52	Antal dygn 53 Antal fraktioner 54
Kritiskt organ I 44 Dos, rad Plan 45 Mitt 46 " " II 47 " 48 49 " " III 50 " 51 52	Antal dygn 53 Antal fraktioner 54 				

Fig. 11 Front page of proforma B.

TABLE VIII. Cont'd.

RADIOTHERAPY - EXTERNAL BEAMS

Start of treatment	Date
Completion of treatment	Date
Planning	No/yes
Series No.	z
Field parameters:	
Field No.	zz
Target	Anatomic code (III)
Level of planning section	Code (III)
Radiation quality	Code
Technique of irradiation	Code
SSD (SAD)	zzz cm
Field size, rectangular	zz x zz cm
" " , circular	zz cm diam
" " , irregular	zzz cm ² eg.sq.
Central axis coordinates	Angle, radius (III)
Beam direction	Angle (III)
Filter(s)	Wedge, orientation (III), bolus etc.
Number of treatments to field	zzz
Field dose	zzzzz rad
Maximum target dose	zzzzz rad (III)
Minimum target dose	zzzzz rad (III)
Representative dose	zzzzz rad (III)
According to plan	*
In vivo dose measurements	*
Number of days	zzz
Number of fractions	zzz
Critical organs:	
Organ	Anatomic code (III)
Dose	zzzzz rad
According to plan	*
In vivo dose measurements	*

YTAPPLIKATION; INTRAKAVITÄR, INTERSTITIELL OCH INTERN STRÅLBEHANDLING

Behandling påbörjad den 1

år	mån	dag

avslutad " 2

--	--	--

Fält Beh.n ^o	Target	Strålkälla	Ant. prep.	Σ	mCi	antal beh	Behandlingstid d h m	Dos i target, rad
3	4	5	6	7	8	9	10	11
3	4	5	6	7	8	9	10	12

CYTOSTATICABEHANDLING

Beh. påbörjad		Beh. avslutad		Typ, administration	Total dos, mg	Biverkningar									
år	mån	dag	år			mån	dag	blod avfall sllh smärtor övr.							
20			21			22		23		24	25	26	27	28	31
20			21			22		23		24	25	26	27	28	32
20			21			22		23		24	25	26	27	28	33

HORMONBEHANDLING

Beh. påbörjad		Beh. avslutad		Typ	inj	per os	Biverkning					
år	mån	dag	år				mån	dag	Nej	Ja		
40			41			42		43	44	45	46	51
40			41			42		43	44	45	46	52

Fig. 12 Back page of proforma B.

TABLE VIII. Cont'd

RADIOTHERAPY - SURFACE APPLICATION, INTERSTITIAL and INTRACAVITARY

Treatment started	Date
Treatment concluded	Date
Field/treatment parameters:	
Field/Treatment No.	zz
Target	Anatomic code (III)
Source of radiation	Code
Number of sources	zzz
Total activity	zzz.z mCi
Number of treatments	zz
Treatment time	zz d zz h zz min
Target dose	zzzzzz rad

RADIOTHERAPY - INTERNAL (METABOLIC)

Treatment given	Date
Treatment No.	zz
Target	Anatomic code (III)
Nuclide	Code
Total activity	zzz.z mCi
Target dose	zzzzzz rad

CHEMOTHERAPY

Treatment started	Date
Treatment concluded	Date
Drug/combination	Code
Route of administration	Code
Total dose	zzzzz.z mg
Side effects	Bone marrow/skin, hair/mucous membranes/nervous system/ other

HORMONAL THERAPY

Treatment started	Date
Treatment concluded	Date
Drug	Code
Route of administration	I.m./oral
Side effects	No/yes

		Formulärnr 04 Pat intagen den 1 <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> " utskriven " 2 <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> Poliklinikbesök 3 <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> Brev el. svarsformulär 4 <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table>	
RADIOLOGISKA KLINIKEN I LUND			
DIAGNOS: 76 <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> PAT: Besvär: 5 Förbättrad 6 St g 7 Föremanad 8 Lever tillst okänt 75 TUMÖR: Läkt 9 Regress 10 St g 11 Progress 12 Lokalt recidiv 13 Ny primärtumör 25		Mors den 14 <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> Obdu- cerad 15 16 Död i sin sjuk 17 På kyrk 18 med metast. 19 Död i interkurr 20 lokalt läkt 21 utan metast. i annan ca. 22 Diagnos: 24 <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table>	
LYMFKÖRTEL- ^{Inga} METASTASER 26 ^{Regress} 27 ^{St g} 28 ^{Progress} 29 ^{Ny tillkomna} 30 ORGAN- ^{Inga} METASTASER 32 ^{Regress} 33 ^{St g} 34 ^{Progress} 35 ^{Ny tillkomna} 36 Lokel: 31 <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> Lokel: 37 <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table>			
STRÅLREAKTION UNDER BEHANDLING HUD: ej märkbar 38 SLEM-HINNOR: ej märkbar 43 rodnad 39 rodnad 44 torr epidermit 40 kärlinj. ödem 45 vätskande 41 fibrin, blodn. 46 nekros 42 nekros 47		STRÅLREAKTION EFTER AVSLUT. BEHANDLING HUD: 48 pigmentering 54 SLEMHINNOR 49 ödem 55 50 atrofi 56 51 teleangiect 57 52 fibros 58 53 nekros 59	
Beh. avbruten p.g. stråleaktion 60 Annan reaktion eller komplikation: 61 <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> + <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table>			
☐ Behandl. planeras el. pågår: operation 62 strålebehandling 63 cytostatika 64 hormoner 65 Ingen ytterligare beh. 66 Pat. skall återkomma den <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> 67 Följes ej 68 Avi till inrem. läkare 69 Pat. remitt. till: <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> 70 Efterhör där om <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> mån. 71 Efterhör patienten om <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> mån. 72 Efterhör PA om <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> mån. 73 Efterhör <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> 74 Datum <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> om <table border="1" style="display: inline-table; width: 100px; height: 20px;"></table> mån.			

Fig. 13 Front page of proforma C.

TABLE VIII. Cont'd.

FOLLOW-UP

Admission to ward	Date
Discharge from ward	Date
Visit to out-patient clinic	Date
Contact by mail	Date
Diagnosis No.	Internal code
Patient's condition	Free from symptoms/regression of symptoms/unchanged/deteriorating/ not known
State of tumour	No evidence of tumour/regression/ unchanged/tumour progression/local recurrence
State of nodes	Non-palpable/regression/unchanged/ advancing/new involvement + anatomic code (III)
State of metastases	No evidence/regression/unchanged/ advancing/new involvement + anatomic code (III)
New primary	*
Reaction during treatment	Skin(5)/mucous membranes(5)/other
Treatment discontinued due to reaction	*
Reaction after concluded treat- ment	Skin(6)/mucous membranes(6)/other
Measures planned	Surgery/radiotherapy/chemotherapy/ no further treatment
Next visit to out-patient dept	Date
Follow-up concluded	*
Pat referred to	Text
Mailing routine inquiry	Patient/registrar's office/other + address Interval (months) or date

DEATH

Date of death	Date
Autopsy performed	No/yes
Cause of death	Current/intercurrent disease
State of tumour	Primary present/not present
State of metastases	Present/not present
Dead in new malignancy	ICD-8 code

TABLE IX. Suggested MUMPS data base.

^ID

YY

- MMDDNNN/100	C (Check digit of CRN)
- - ID	
- - - 1	Name
- - - 2	Parish
- - - 3	Occupation/Marital state
- - - 4	Address
- - DIA	No. of diagnoses
- - - DX#	ICD/Stage/TNM/Rad No./Revised Dx/Text
- - - DX#+0.01	SNOP/Type/Date/Specimen No./Department/Text
- - - DX#+0.02	Date of referral/Department/Address
- - - DX#+0.03	Date of admission
- - - DX#+0.04	Site of primary/nodes/metastases
- - - DX#+0.05	Cause of death/State of primary/metastases
- - - DX#+0.06	Date/Present status
- - - DX#+0.07	Not accepted for treatm./Ca.reg./Not included
- - MOR	Date of death/Autopsy
- - EXT	(External beam therapy)
- - - 0	No. of treatments
- - - #/100	Dx No./Date of start/Radiation quality/Series No.
- - INT	(Intracav. or interstitial therapy)
- - - 0	No. of treatments
- - - #/100	Dx No./Date of start/Type of source/Activity
- - MET	(Metabolic therapy)
- - - 0	No. of treatments
- - - #/100	Dx No./Date/Administration/Nuclide/Activity
- - KEM	(Chemotherapy)
- - - 0	No. of treatments
- - - #/100	Dx No./Date of start/Drug or combination
- - OP	(Surgery)
- - - 0	No. of operations
- - - #/100	Dx No./Date/Type/Department/Radicality

TABLE IX. Cont'd.

^EXT

YY

- MMDDNNN/100	C (Check digit of CRN)
- - 0	No. of treatments
- - BEH#	Diagnosis No.
- - BEH#+0.01	No. of subrecords (SUB)
- - - SUB#x10	
- - - SUB#x10+ $\frac{IV\#}{100}$	Value (No. and value of variable IV as defined in I).

^FU

YY

- MMDDNNN/100	C (Check digit of CRN)
- - 0	Present status/date of conclusion of follow-up
- - 1 (PFU)	No. of changes/No. of visits
- - 1+#/100	Date/Status/Measure/Date of next visit
- - 2 (SFU)	No. of changes/No. of mail contacts
- - 2+#/100	Date/Status/Measure/Department/Interval

REFERENCES

- ARNESJÖ B., GÖRANSSON G. and MÖLLER T.R.: Regionalt tumörregister - ett förslag till organisation och funktion (In Swedish). *Läkartidningen* 70 (1973), 4041-4044.
- BARCLAY T.H.C.: The evaluation of data collection procedures - Assessment of completeness. In: *Cancer Registry. Recent Results in Cancer Research* 50, Springer-Verlag, Berlin 1975, p.59-75.
- BLEEHEEN N.M., MacKINNON D.M., HARRIS F.T.C. and BUCKLEY-SHARP M.D.: An analysis of error in the retrieval of patient case note data. *J. Laryngol.* 84 (1970), 689-695.
- BODOR F., DAHLERUS B., GUSTAFSSON T., HÅKANSSON C.-H. and LEVAN A.: A method for the mathematical study of the stained mass of the chromosomes. *Hereditas* 77 (1974), 63-72.
- BOLANDER A.-M.: Linkage of census and death records to obtain mortality registers for epidemiological studies in Sweden. *Proc. XI International Cancer Congress, Florence 1974. Excerpta Medica, Amsterdam 1975*, vol. 3, p. 36-39.
- BOWIE J. and BARNETT G.O.: MUMPS - An economical and efficient time-sharing system for information management. *Computer Programs in Biomedicine* 6 (1976), 11-22.
- BRINCKER H.: Medicinsk og administrativ databehandling ved Radiumstationen i Odense (In Danish). *Ugeskr. Laeg.* 133 (1971), 612-618.
- BRINCKER H.: The design and implementation of an EDP-system in clinical routine - The collection of clinical data. In: *Computers in Radiotherapy in Europe. Proceedings of a Workshop, Vienna 1976. British Institute of Radiology, London 1977. Special report 13* (In press).
- BROSS I.D.J., GREGORY A.R., RATHOD C. and BROSS D.S.: Unobtrusive medical data-input systems. *Int. J. Bio-Medical Computing* 4 (1973), 219-228.
- BRUCE W.R., JENKIN R.D.T. and GORDON S.: A physician-oriented medical index for hospital or clinic use. *Can. Med. Ass. J.* 102 (1970), 1080-1084.
- BRUCE W.R.: Application of optical character recognition and microfiche to medical records. *Annual Report, The Ontario Cancer Treatment and Research Foundation, 1971.*

- COLLEN M.F. (Ed.): Hospital computer systems. John Wiley & Sons, New York 1974.
- DAHLIN H.: A survey of computers and computer programs in radiotherapy. Proc. Fourth Int. Conf. on the Use of Computers in Radiation Therapy Uppsala 1972, p. 174-235.
- DEELEY T.J.: Computers in Radiotherapy - Clinical Aspects. Butterworth, London 1972.
- DUTREIX J., DUTREIX A. and ZUMMER K.: Data recording in radiotherapy. Computer Programs in Biomedicine 2 (1972), 232-241.
- EDWARDS D.N.: The radiotherapeutic record. In: Computers In Radiology. Proc. Int. Meeting, Brussels 1969. Karger, Basel 1970, p.217-220.
- EDWARDS D.N.: Clinical records and electronic data processing. Clin. Radiol. 23 (1972), 117-126.
- EDWARDS D.N.: The standard clinical record. In: Radiology vol. 2, Proc. XIII Int. Congress of Radiology, Madrid 1973. Excerpta Medica. Amsterdam 1974, p. 680-684.
- EDWARDS D.N.: Do we need a computer? Clin. Radiol. 25 (1974), 19-25.
- ELLIS F.: Computers and planning in cancer treatment. In: Role of Computers in Radiotherapy. Report of a Panel, Vienna 1967. International Atomic Energy Agency, Vienna 1968, p 145-161.
- ELLIS F. and PAINE C.H.: Rapid and accurate recording of cancer data for computer analysis. In: Computers in Radiology. Proc. Int. Meeting, Brussels 1969. Karger, Basel 1970, p 233-236.
- FAYOS J.V.: Computerized radiation therapy data. Cancer 37 (1976), 2736-2745.
- FEINSTEIN A.R.: Quality of data in the medical record. Comput. Biomed. Res. 3 (1970), 426-435.
- FENICHEL R.R. and BARNETT G.O.: An application-independent sybsystem for free-text analysis. Comput. Biomed. Res. 9 (1976), 159-167.
- FOLLAND R.E., NEURATH P.W. and POWELL W.T.: The Tufts-New England Medical Center computer-based tumor registry. In: Computer Applications in Radiation Oncology, Proc. Fifth Int. Conf. on the Use of Computers in Radiation Therapy, Hanover 1974. University Press of New England, Hanover 1976, p 481-489.
- FONDARAI J.A.: Statistical use of data in radiotherapy. In: Radiology. vol 2, - Proc. XIII Int. Congress of Radiology, Madrid 1973. Excerpta Medica, Amsterdam 1974, p 689-692.
- FREY H.S.: An on-line graphic inquiry system for analysis of tumor registry data. In: Computers in Biomedical Research, vol. III, Academic Press, New York 1969, p. 87-105.

- GARY-BOBO J. and ROSSI M.: Premier bilan de 3 années d'utilisation de l'ordinateur IBM 360-40 pour l'exploitation des dossiers du Centre régional de Lutte contre le Cancer de Montpellier. In: Computers in Radiology. Proc. Int. Meeting, Brussels 1969. Karger, Basel 1970, p. 237-240.
- GELL G. and SCHWARZ G.: AURA: Automatische Befunddokumentation durch Klartextverarbeitung. Fortschr. Röntgenstr. 121 (1974), 384-388.
- GRAEPEL P.H., HENSON D.E. and PRATT A.W.: Comments on the use of the Systematized Nomenclature of Pathology. Meth. Inform. Med. 14 (1975), 72-75.
- GREENWOOD R.M.: KODIAC, a system for disease coding by a medium-sized computer. Int. J. Bio-medical Computing 3 (1972), 123-134.
- GROS Ch., VROUSOS C., BLÉGER A. and MARTIN J.: Le dossier sénologique. In: Computers in Radiology. Proc. Int. Meeting, Brussels 1969. Karger, Basel 1970, p. 248-251.
- GUSTAFSSON T.: A new hard-copy colour display system using ink jets and its application in nuclear medicine. In: Information Processing in Scintigraphy. Proc. Third Int. Conf. on Data Handling and Image Processing in Scintigraphy, Cambridge, Mass. 1973. USERDA Technical Information Center, Springfield 1973, p. 195-203.
- HALL P., MELLNER C. and DANIELSSON T.: J 5 - A data processing system for medical information. Meth. Inform. Med. 6 (1967), 1-6.
- HENDRICKSON L. and MYERS J.: Some sources and potential consequences of errors in medical data recording. Meth. Inform. Med. 12 (1973), 38-45.
- HUTCHISON G.B.: Criteria of cure: Statistical considerations. Natl. Cancer Inst. Monogr. 36 (1973), 561-565.
- ICRU Report 24: Determination of Absorbed Dose in a Patient Irradiated by Beams of X or Gamma Rays in Radiotherapy Procedures. International Commission on Radiation Units and Measurements, Washington 1976.
- ICRU Report: Dose Specification for Reporting External Beam Therapy. (In preparation).
- International Classification of Diseases, Injuries and Causes of Death, 1965 revision. World Health Organization, Geneva 1965.
- JERN M.: Color Jet Software Manual (version V). Lund University Computer Centre 1976.
- KAHR E. and MOSKON A.: Medizinische Dokumentation und Mahnverfahren bei krebserkrankten Bestrahlungspatienten. Strahlentherapie 127 (1965), 382-392.
- KENNEDY A.D. and GEHAN E.A.: Computerized simple regression methods for survival time studies. Computer Programs in Biomedicine 1 (1971), 235-244.

- KOREIN J., TICK L.J., WOODBURY M.A., CADY L.D., GOODGOLD A.L. and RANDT C.T.: Computer processing of medical data by variable-field-length format. *JAMA* (1963), 132-138.
- KOREIN J.: The computerized medical record: The variable-field-length format system and its applications. In: *Information Processing of Medical Records. Proc. IFIP-TC 4 Working Confer. on Informtion Processing of Medical Records*, Lyon 1970. North-Holland Publ. Comp. Amsterdam 1970, p. 259-291.
- KRICHEFF I.I., CHASE N.E. and KOREIN J.: Computer processing of clinical medical data. In: *Computers in Radiology. Proc. Int. Meeting*, Brussels 1969. Karger, Basel 1970, p. 289-316.
- KÖHLER C.O., WAGNER G. and WOLBER U.: Interactive data processing in medicine - Man-machine dialogue (Bibliography). *Meth. Inform. Med.* 15 (1976), 102-120.
- LANE F.W. (Chairman), CLAMPITT S., COHEN L., DAVIS L., GRIEM M. L., HENRY J.A., OLSEN B., PREWITT J.M.S., RAGAN D.P. and STOVALL M.: Reports of workshop directors on the establishment of an international radiotherapy data bank. In: *Computer Applications in Radiation Oncology. Proc. Fifth Int. Conf. on the Use of Computers in Radiation Therapy*, Hanover 1974. University Press of New England, Hanover 1976, p. 335-372.
- LASZLO J., COX E. and ANGLE C.: The hospital tumour registry. *Cancer* 38 (1976), 395-402.
- LINDGREN M.: Organisation der Nachbehandlung und Kontrolle bestrahlter Geschwulstkranker. *Der Radiologe* 2 (1962), 309-312.
- MacKINNON D.M., BLEEHEN N.M., BUCKLEY-SHARP M.D., HARRIS F.T.C. and BRACEWELL A.: A record system for carcinoma of the larynx and hypopharynx designed for computer analysis. *J. Laryngol.* 84 (1970), 679-688.
- Manual of Tumor Nomenclature and Coding, 1968 ed. American Cancer Society 1968.
- MARTIN J.M., POINTEL J.P., MARTIN J. and DEBRY G.: Structural, syntactical and semantic data validation in the computer processing of a medical file. *Meth. Inform. Med.* 15 (1976), 205-210.
- McAINSH T.F., ORR J.S., HALNAN K.E. and STEVENSON J.: Automated data processing of a radiotherapy case record. *Int. J. Bio-Medical Computing* 2 (1971), 137-145.
- McLEAN E.R., FOOTE S.V. and WAGNER G.: The collection and processing of medical history data (Bibliography). *Meth. Inform. Med.* 14 (1975), 150-163.

- MELLNER C., SELANDER H. and WOŁODARSKI J.: The Karolinska Hospital information system. *Meth. Inform. Med.* 13 (1974), 125-140.
- MESEL E., WIRTSCHAFTER D.D., CARPENTER J.T., DURANT J.R., HENKE C. and GRAY E.A.: Clinical algorithms for cancer chemotherapy - systems for community-based consultant-extendors and oncology centers. *Meth. Inform. Med.* 15 (1976), 168-173.
- MINET P. and TIBAUX G.: Premières impressions au sujet du traitement par l'ordinateur des dossiers de patients cancéreux. In: *Computers in Radiology. Proc. Int. Meeting, Brussels 1969.* Karger, Basel 1970, p. 241-247.
- MINET P.: Creation of a data bank. In: *Radiology, vol. 2. Proc. XIII Int. Congress of Radiology, Madrid 1973.* Excerpta Medica, Amsterdam 1974, p. 687-688.
- MUG-3 Master List of Mumps Applications and Institutions. Document prepared by Mumps Users' Group, St. Louis, 27 May 1976.
- MÖLLER T.R.: ADB-system för produktionskontroll och follow-up (In Swedish). *Nordisk Medicin* 83 (1970), 730.
- MÖLLER T.R., GUSTAFSSON T., NAVERSTEN Y. and CEDERQUIST E.: Experiences of computer-evaluation of thyroid scans in clinical routine. In: *Medical Radioisotope Scintigraphy II. Proc. of a Symposium, Monte Carlo 1972.* International Atomic Energy Agency, Vienna 1973, p. 519-527.
- MÖLLER T.R.: Documentation of external beam therapy. In: *Computers in Radiotherapy in Europe. Proc. of a Workshop, Vienna 1976.* British Institute of Radiology, London 1977, Special report 13. (In press).
- NEURATH P.W., BLOEDORN F.G., MUNZENRIDER J.E. and PEDRAZZI M.G.: A selective patient information system. *Cancer* 33 (1974), 1653-1663.
- OKUBO R.S., RUSSELL W.S., DIMSDALE B. and LAMSON B.C.: Natural language storage and retrieval of medical diagnostic information. *Computer Programs in Biomedicine* 75 (1975), 105-130.
- OLSON K.B., STONE M. and HARLAN A.: The hospital tumor registry. *Pathology Annual* 5 (1970), 257-277.
- PARZIALE V.S. and KAPLAN H.S.: Information storage and retrieval system for the study of Hodgkin's disease. *Natl. Cancer Inst. Monogr.* 36 (1973), 541-553.
- PHILLIPS T.L.: Personal communication (1972).
- PREWITT J.M.S.: Clinical motivations for an international radiotherapy data base (Abstract). In: *Computer Applications in Radiation Oncology. Proc.*

- Fifth Int. Conf. on the Use of Computers in Radiation Therapy, Hanover 1974. University Press of New England, Hanover 1976, p. 532-533.
- RADATA - Koordineret indsamling og bearbejdning af kliniske data for radioterapeutiske patienter (In Danish). Report circulated to members of Dansk Selskab for Radioterapi og Cancerologi, October 1972.
- ROBINSON III R.E.: Acquisition and analysis of narrative medical record data. *Comput. Biomed. Res.* 3 (1970), 495-509.
- SCHWAB W.: Aktuelle Bemerkungen zur Anwendung des TNM-Systems im Kopf-Hals-Bereich. *Laryng. Rhinol.* 54 (1975), 44-64.
- SCUDIERO C.A. and WONG R.L.: A numerically coded file of operative procedures derived from a free text data collection system: A measure of the accuracy. *Meth. Inform. Med.* 15 (1976), 21-28.
- SNELLING M.D. and KUNKLER P.B.: A computer programme for the records of patients treated by radiotherapy for any form of malignant disease. In: *Computers in Radiology. Proc. Int. Meeting, Brussels 1969.* Karger, Basel 1970, p. 221-227.
- SNELLING M.D.: The radiotherapeutic record - Report of the round table. In: *Computers in Radiology. Proc. Int. Meeting, Brussels 1969.* Karger, Basel 1970, p. 252-265.
- SPIESSL B., DOLD U. and OTT G.: Proposal to add a factor of evidence to the TNM system (Letter to the editor). *Int. J. Cancer* 7 (1971), 372-373.
- SPRI råd 4.7: Grundjournal (SPRI advice 4.7: Basic medical record). Sjukvårdens och Socialvårdens Planerings- och Rationaliseringsinstitut, Stockholm 1976.
- STEBEN J.D., WU R., SUNTHARALINGAM N. and MANSFIELD C.: The Thomas Jefferson Hospital radiation therapy departmental computer system. In: *Computer Applications in Radiation Oncology. Proc. Fifth Int. Conf. on the Use of Computers in Radiation Therapy, Hanover 1974.* University Press of New England, Hanover 1976, p. 399-402.
- STRÜTER H-D. and HALAMA J.: Ein Organisationssystem in der Strahlentherapie mit Hilfe einer institutseigenen EDV-Anlage. I - Patientendaten und Datenfluss, II - Systemorganisation und Informationsausgabe. *Strahlentherapie* 148 (1974), 259-270 (I), 271-278 (II).
- SUNDSTRÖM A.: Personal communications (1970-1975).
- Systematized Nomenclature of Pathology. College of American Pathologists 1965.
- The Second Report of the Joint Working Party on Computers in Radiotherapy. *Clin. Radiol.* 25 (1974), 167-173.
- TSIEN K.C.: The application of automatic computing machines to radiation treatment planning. *Br. J. Radiol.* 28 (1955), 432.

- TULINIUS H.: The ninth revision of ICD chapter on neoplasms and its application in oncology. In: Cancer Registry. Recent Results in Cancer Research 50, Springer-Verlag, Berlin 1975, p. 33-37.
- UICC: TNM Classification of malignant tumours. International Union Against Cancer, Geneva 1974.
- UICC: TNM General rules. International Union Against Cancer, Geneva 1974.
- WAGNER G.: Quality control and error checking. In: Information Processing of Medical Records. Proc. IFIP-TC 4 Working Conf. on Information Processing of Medical Records Lyon, 1970. North-Holland Publ. Comp., Amsterdam 1970, p. 222-236.
- WAKEFIELD J.S. and YARNALL S.R.: Implementing the problem-oriented medical record, 2nd ed. Medical Computer Services Association, Seattle 1976.
- WEED L.L.: Medical records that guide and teach. N.Engl. J. Med. 278 (1968), 593-600.
- WELLS A.H.: Systematized Nomenclature of Pathology. Minnesota Medicine 55 (1972), 585-590.
- WHEELER T.K., RUBERY E.D. and HAYBITTLE J.L.: A computer based clinical record system using mark-sense forms. Br. J. Radiol. 49 (1976), 863-867.
- WHO Handbook for standardized cancer registries. Offset publ. No. 25. World Health Organization, Geneva 1976.
- WICKHAM J.E.A., CHARLTON C.A.C., RICHARDS B., HENDRY W.F., WARD J.P., O'DONOGHUE E.P.N., HAMSHERE R.G. and FRANKLIN D.A.: A computer-based record and organisation system for a department of urology. Br. J. Urol. 47 (1975), 345-357.

